Bilingual advantages in middle-aged and elderly populations

Houtzager, Nienke

Chapter 3

Aging and bilingual processing: Age-related and individual differences in groups of early bilingual Frisians.

Abstract

We conducted a study to investigate effects of bilingualism on nonverbal and verbal cognitive performance. In this chapter we report on the results from the first part of our study, in which 26 middle-aged (MA 46) and 26 elderly (MA 73.2) early bilingual speakers of Dutch and Frisian performed a verbal fluency test, a task switching test and a working memory (WM) test. In an analysis focusing on between- and within-group differences, we found that the age-groups differed significantly in the task switching and WM tests and in the category switching component of the verbal fluency task. Analyzing the effects of other factors, we found that the higher the degree of balance between participants' languages, the smaller their costs for switching between tasks.

This chapter has been published as a research paper:

Houtzager, N., Lowie, W.M. & De Bot, K. (2014). Bilingualism, aging and multilingual processing: Age-related and individual differences in groups of early bilingual Frisians. In L. Filipović and M. Pütz (eds.), *Multilingual Cognition and Language Use*, pp 263-285. Amsterdam: Benjamins.

We thank the anonymous reviewers who evaluated this manuscript for their questions, feedback and suggestions on previous versions of this work, and Mik van Es for his invaluable help and support with the statistical analyses.

3.1. Introduction

During the past two decades, the general attitude towards bilingualism has changed, the focus having shifted from concerns about harmful effects of bilingualism for children to the practical benefits of fluency in multiple languages. Moreover, there has been evidence suggesting that bilingualism may positively affect general cognitive performance across the lifespan: bilinguals were found to outperform monolinguals in nonverbal tasks depending on executive control, an advantage that was found both in children (c.f. Bialystok, 2001) and in adults (Costa, Hernández & Sebastián-Gallés 2008), although not all studies report significant group differences. On the other hand, linguistically, bilinguals seem to be at a disadvantage compared to monolinguals: bilingual children control smaller vocabularies in each of their languages than monolingual children (Mahon & Crutchley, 2006), and differences in verbal performance continue into adulthood and old age, when bilinguals exhibit deficits in vocabulary access (see Kaushanskaya & Marian, 2007) and lexical decision (Ransdell & Fischler, 1987), and are also reported to show more tip of the tongue experiences than monolinguals (cf. Gollan & Acenas, 2004).

We have started a large-scale study to investigate effects of bilingualism on cognitive performance. With this aim, we compare performance on a number of cognitive tasks between adult groups of early bilinguals, late bilinguals and functional monolinguals. First, we look at effects of bilingualism on linguistic performance. Second, we investigate bilingual effects on executive control functions that may play a part in language processing. In view of the interaction between bilingualism and aging in these effects, we compare the performance of elderly and middle-aged participants, both in verbal and nonverbal tasks. Third, we address the hypothesis that bilinguals' constant conflict between competing language systems can simultaneously result in deficits in lexical access and in a boost of executive control functions, because of their constant involvement in multiple language processing (Bialystok, Craik & Luk, 2008a; Bialystok, 2009). With this in mind, we analyze the interaction between the results of the verbal and nonverbal tasks of our experiment.

In this chapter, we report on the results from the first part of our study, in which we tested middle-aged and elderly early bilingual speakers of Dutch and Frisian. Frisian is a minority language spoken in Friesland, a northern province of the Netherlands, and is closely related to Dutch. For our comparison with the other two language groups, we needed a homogeneous group of bilinguals, and because our participants were all balanced bilinguals in the same two languages, we expected our group to fit this criterion of homogeneity. The main research question of this chapter concerns the issue to what extent individual differences in language background and language use play a role in task performance within this group of early bilinguals.

In the first section of this chapter, we present an overview of research that has been done in the field of verbal production and verbal fluency. In the first place, this leads to a justification of the choice of the verbal task used in our study. Secondly, we aim to investigate which factors, apart from language-related variables, can be expected to predict performance on our tasks, on the basis of results from earlier research. Thirdly, we want to give a brief overview of earlier findings on bilingual performance on verbal fluency tasks, so that we can see how it may differ from monolingual performance. This part is followed by a short section on bilingualism and executive control. Here as well we want to explain the choice of our cognitive task, and second, we want to investigate to what extent both language-related and other factors can be expected to affect task performance. Next, we show how we define and operationalize a number of factors concerning language use and language background. This section is followed by a description of the design and procedure of our experiment, and finally by a presentation of the first results and a brief discussion of our findings.

3.2. Verbal production and verbal fluency

In spite of the apparent ease most bilinguals reveal in using their languages and in switching between them, being bilingual also implies considerable language processing costs, which seem related to differences in vocabulary size. Bilingual children are consistently found to have smaller vocabularies than monolingual children (Mahon & Crutchley, 2006). Although such a difference does not necessarily persist in adults (cf. Portocarrero, Burright & Donovick, 2007; Fernandes, Craik, Bialystok & Kreuger, 2007), bilingual adults are often reported to be slower in lexical access. This is particularly reflected in lexical production tasks asking for a timely response, such as picture naming, in which bilinguals proved consistently slower, even when tested in their first and dominant language (see Ivanova & Costa, 2008), and verbal fluency tests. The verbal fluency test is now often used in research on bilingual vs. monolingual language processing. It is a free-naming task, in which participants, within one minute, name as many exemplars as possible from a number of semantic or phonological categories. Thus, the test involves the activation of large semantic networks and as such, of many related lexical representations. According to Sandoval, Gollan, Ferreira and Salmon (2010), this makes it a valuable assessment tool for analyzing potential effects of between-language interference in bilinguals.

Another reason why the verbal fluency test fits our purpose of analyzing bilingual verbal performance is that the test consists of two different components – semantic and phonological, or letter fluency - which are generally considered to tap into different cognitive processes (Troyer, Moscovitch & Winocur, 1997, Roselli et al., 2000; Gollan et al., 2002; but see Unsworth, Spillers & Brewer, 2011, for counterevidence: they found no systematic differences between semantic and letter fluency tasks, in terms of total number of items generated, cluster size, or frequency of switches). During the semantic fluency task, participants are usually required to name exclusively nouns, particularly of a concrete nature. Consequently, all the time items are activated within that particular semantic field; this determines their accessibility, so that performance on the task depends to a large extent on the inherent organization of the participant's semantic knowledge (see Portocarrero, Burright & Donovick 2007; Sauzéon et al., 2011). During the phonological fluency task, on the other hand, words cannot be named related to their meaning (Perret, 1974). This makes the task more difficult for most participants, because they have to inhibit the tendency to retrieve words in a similar way as during natural speech. Besides, the search for items to be named during phonological fluency is not restricted to (concrete) nouns, but lexical items from all word classes are possible candidates. This implies that the task taps into more effortful, strategic search processes than semantic fluency. This strong dependence of phonological fluency on inhibitory processes and effortful search strategies explains why it has been argued that performance on this verbal fluency component taps most strongly into executive control functions (Filippetti & Allegri, 2011). This does not mean, however, that performance on semantic fluency is not to some degree dependent on executive control functions as well; stopping a search process and initiating a new one, for instance, is an aspect of the task that taps into executive control functions, too. Sauzéon et al. (2011)

conducted a study involving a verbal fluency test to investigate whether verbal knowledge can compensate for the age-related decline that is often observed in performance on this task. To this end, a dissociation was made between semantic and non-semantic components of the verbal fluency test. Elderly participants indeed appeared to compensate for the age-related decline in task performance, reflected in a positive correlation between age and vocabulary knowledge, but only on phonological fluency. According to Sauzéon and colleagues, the non-semantic component of the task, which is sensitive to age-related decline, includes "speed processing and executive functions, such as strategic search processes, shifting, monitoring, inhibition processes, and working memory" (p. 145). This also explains why it is important to test participants' working memory, and include the results on this task in the analyses of the verbal fluency results.

In our verbal fluency test, we also included the component category switching. During this task, participants have to switch between naming items from two different semantic categories, for instance animals and jobs. This task component taps to a large extent into the same cognitive processes as semantic fluency, but the element of having to switch between two different categories imposes an extra demand on executive control functions, in particular the function that Miyake et al. (2000) defined as mental flexibility or switching between different mental sets, and working memory, defined by the same authors as the executive function of updating information. One of the few studies that also included the switching condition in the test battery is an experiment by Mayr and Kliegl (2000), involving young (MA 23.4) and old (MA 67.7) adults. They wanted to investigate whether age-related differences in semantic retrieval tasks are due to semantic processes or to other, task-specific, non-semantic processes. Troyer et al. (1997) had found that an age-related decline in performance on semantic fluency was related to a decline in older participants to switch between semantic clusters. This was considered to reflect an age-related decline in the executive function of switching between mental sets. On the basis of this model, age-differences would be more pronounced in the switch condition than in the non-switch condition of semantic fluency. The other option, proposed by Mayr and Kliegl, is that executive functions are "relevant for every single act of retrieval" (p. 31), implying that age-related differences would also arise in the non-switch condition, and not be much higher in the switch condition. This latter view was supported by their findings, as age-related deficits in set-switching appeared to be relatively subtle.

Finally, the third reason why we chose the verbal fluency test for our experiment is that previous research (cf. Sandoval et al., 2010; Luo, Luk & Bialystok, 2010) has shown that this test also lends itself for a number of qualitative analyses of the responses, such as clustering and switching and time-course analyses. Some of these analyses may enable a more detailed investigation of the mechanisms that are responsible for the differences between bilingual and monolingual verbal production. These analyses are currently worked on and will be reported in a future article.

The second aim of this section concerned the question to what extent individual differences within our target group, besides language-related factors, play a role in task performance. Various studies found large effects of age and level of education on verbal fluency. Acevedo et al. (2000), in a large-scale study on semantic fluency in monolingual English and Spanish participants over age 50, found the most important factors to be age and education. Increasing age was associated with lower fluency scores in all categories (animals, vegetables and fruits), and increasing level of education was associated with higher total fluency scores. They also found that different fluency measures were affected to a different degree by different factors. For instance, scores on vegetable and fruit fluency were associated with age, education and gender (with women outperforming men), but animal fluency only with age and education. This implies that it is very important to include a sufficient number of different semantic categories in the verbal fluency test. Troyer (2000), in a normative study, meant to provide data for clinicians and researchers to determine the nature of fluency impairments, involving 411 healthy English-speaking adults between 18 and 91 years old (MA 59.8, SD 20.7), also found increasing age associated with lower scores, both on semantic and phonological fluency; it was a stronger predictor for semantic than for phonological fluency. In an earlier study (1997), involving 54 older (MA 73.3, SD 6.5) and 41 younger adults (MA 22.3, SD 3.8), Troyer and her colleagues did not find a significant age difference on phonological fluency, but the younger participants significantly outperformed the elderly group on semantic fluency. However, Ostrosky, Ardila and Rosselli (1999), reporting on a large-scale study involving 800 monolingual Spanish speaking participants between 16 and 85 years old, divided over 4 age-groups, found the effect for age stronger for the phonological than for the semantic condition. The effect for education that they found was strongest for performance on phonological fluency, as well. Ardila, Ostrosky-Solis, Rosselli and Gomez (2000) also found the effect of education to be stronger in the phonological than in the semantic condition (accounting for 38.5% vs. 23.6% of the variance). Rosselli and colleagues (2009) conducted an experiment on a group of 105 Hispanic adults aged 55-98, divided into three educational groups, but did not include phonological fluency. They found that increasing levels of education only predicted higher scores on the categories animals and clothing, but not on the categories fruits and vegetables

In summary, age and education can be said to be the main factors besides language-related variables that seem to affect performance on verbal fluency tests. Results differed as to whether the effect was strongest in semantic or phonological fluency, but the tendency was always that lower age and higher level of education are associated with higher fluency scores. In our analyses of the verbal fluency results of our early bilingual group, we will therefore take the variables age and education also into account.

Recently, verbal fluency tests have become widely used in research on differences between monolinguals and bilinguals, as well. What most studies agree on is that generally monolinguals outperform bilinguals on these tasks (e.g. Gollan et al., 2002; Bialystok, Craik & Luk, 2008b); not only do they produce more correct items, but they also tend to show faster first response times (Sandoval et al., 2010). It is clear from these findings that this bilingual disadvantage generally reflects slower lexical access. However, in order to better understand the reasons for this bilingual deficit, we will have a closer look at performance on the different task components, because, as we argued above, these are considered to tap into different cognitive processes. A number of previous studies found that differences between monolinguals and bilinguals were only present in the semantic component of the test. For instance, Rosselli et al. (2002) conducted a study involving 82 elderly participants (MA 61.76, SD 9.3), of which 45 were English monolinguals, 18 were Spanish monolinguals, and 19 Spanish-English bilinguals; the bilinguals were tested both in English and Spanish. The bilinguals produced almost the same number of items on the phonological components of the task, but produced significantly fewer items on one semantic category in Spanish and on both semantic categories in English. Portocarrero, Burright and Donovick (2007) compared 39 monolingual and 39 bilingual college students, and also found that they performed similarly on phonological fluency, but that the monolinguals were significantly better at semantic fluency. Gollan, Montoya and Werner (2002) also conducted a study involving monolingual and bilingual college-aged students, who had to perform a verbal fluency test consisting of 12 semantic, 10 phonological and 2 proper name categories. The monolinguals significantly outperformed the bilinguals on all tasks, but the difference was found to be largest in the semantic fluency components. The difference could not be attributed to language dominance (which was assessed on the basis of self-ratings, by means of a language questionnaire). Rosselli et al. (2000) proposed that this discrepancy between performance on semantic and phonological fluency might be due to the fact that semantic fluency predominantly involves the naming of concrete words, which share more representations across languages than non-concrete words. Therefore, cross-language interference would be stronger in semantic than in phonological fluency.

On the other hand, Obler, Albert and Lozdwick (1986), comparing groups of monolingual and bilingual elderly (MA 73) participants, found the bilinguals outperformed the monolinguals in animal list generation. However, the bilingual group here only counted 11 participants. Bialystok, Craik and Luk (2008b) report on a study in which a group of 24 monolingual English college students (MA 20.7) outperformed a group of bilingual compeers (MA 19.7) in phonological, but not in semantic fluency. Because in previous studies bilinguals were often found to have a smaller vocabulary size (cf. Fernandes, Craik, Bialystok & Kreuger, 2007) the researchers decided to control for vocabulary size, and interestingly, under this condition the bilingual disadvantage disappeared. A second study, reported on in the same article, involved a group of 16 monolingual and 50 bilingual college students; on the basis of results on the Peabody Picture Vocabulary test, the latter group was divided into a group with a high vocabulary (HP) and one with a low vocabulary size (LP). Moreover, in this study the letter fluency task was designed in such a way that it used more restrictions on the items that could be named, and as such could be expected to place higher demands on executive control functions. On semantic fluency, there was no significant difference between the monolinguals and the HP bilinguals, and both groups outperformed the LP bilinguals. On phonological fluency, there was no significant difference between monolinguals and LP bilinguals, but both groups were outperformed by the HP bilinguals. The authors explain the results by arguing that phonological fluency depends on executive processes to monitor and control the interference of irrelevant competitors; because these executive processes are enhanced in bilinguals, bilinguals perform better in lexical retrieval on this task than monolinguals, providing that vocabulary size is controlled for. In a second article (Bialystok, Craik & Luk, 2008a), the authors also involved a group of 24 monolingual (MA 67.2) and 24 bilingual (MA 68.3) elderly participants. Here, the monolinguals outperformed the bilinguals on all tasks, but there was no interaction between age and language group. Sandoval et al. (2010) reported on a study involving 30 English-speaking monolinguals (MA 19.67) and 30 Spanish-English bilinguals (MA 20.33) with similar educational backgrounds; in a second study, 45 bilinguals participated. In the first study, bilinguals showed fewer correct responses in 22 out of 24 phonological categories, and in 11 out of the 15 semantic categories.

In sum, in verbal fluency tests bilinguals were generally found to be at a disadvantage compared to monolinguals. Differences between these groups were found particularly in performance on semantic fluency, which is usually considered to be most dependent on lexical access. A few studies found the largest group difference in performance on phonological fluency, which is considered to be more dependent on executive control functions. However, this bilingual disadvantage disappeared once vocabulary size was controlled for, so that in one study a group of high proficient bilinguals even outperformed the monolingual group.

3. 3. Bilingualism and executive control.

The first evidence for bilingual advantages in general cognitive performance was found in children. Bilingual children were found to outperform monolingual children in metalinguistic tasks (Bialystok, 1988), and later evidence was found that this bilingual advantage might extend to nonverbal domains as well (cf. Bialystok & Martin, 2004; Carlson & Meltzoff, 2008). The explanation proposed for these findings was that bilingual children might develop a mechanism that controls attention to the target-language, and that inhibits the non-target language from interfering; this mechanism might benefit bilinguals in other, non-linguistic, domains as well (see Bialystok, Craik, Klein & Viswanathan, 2004). When research on bilingual effects on cognitive functioning was extended to adults, a bilingual advantage was found in the Simon task (Bialystok et al., 2004), in the anti-saccade task (Bialystok, Craik & Ryan, 2006) and in the ANT task (Costa, Hernández & Sebastián-Gallés, 2008), and bilinguals were also found to show a larger attentional blink, suggesting stronger reactive inhibition of irrelevant information (Colzato et al., 2008). Moreover, the bilingual advantage reported by Bialystok and colleagues (2004) increased in groups of elderly participants (age>60), suggesting that bilingualism might lead to a modulation of age-related effects. However, there has been no agreement as to what mechanism is responsible for these bilingual advantages. Still, a considerable number of studies failed to find differences in performance between bilinguals and monolinguals. For instance, Meuter and Orr (lecture ISB8 Oslo, 2011), in a study involving younger and elderly monolinguals and bilinguals with various language and educational histories, found that the factor bilingualism did not have any impact on performance on a set of different executive function tasks; factors that did have an impact were increased L2 proficiency and level of education. In experiments involving young adults results are mixed, the explanation being that this age-group might be at a peak as regards executive control abilities, so that ceiling effects prevent the detection of group differences (see Costa et al., 2008, Prior & MacWhinney, 2010). The results from a number of experiments involving older adults, such as the reduced Simon effect in bilinguals found by Bialystok et al. (2004) have not yet been replicated. This inconsistency in results is discussed in an overview of studies comparing bilinguals and monolinguals in tasks requiring attentional control (i.e. either conflict resolution or response suppression) by Costa and colleagues (2009). The authors relate a potential bilingual impact on the magnitude of the conflict effect to "the need of bilinguals to continuously control their two languages, focusing on the relevant linguistic representation while avoiding interference from the non-intended language" (p. 136); the impact on overall RTs, however, is related to monitoring processes, which are needed for the effective implementation of these conflict resolution processes, and thus to the process of "settling the language in which communication will proceed" (p. 143), in other words, to monitoring "which language to produce in each communicative interaction" (p. 144). In sociolinguistic situations where a bilingual's languages are linguistically similar, making it potentially more difficult to keep them apart, bilingual effects on executive control might be most likely to occur. This effect might be stronger when these languages are also used in similar contexts, so that speakers constantly have to monitor their language use. The linguistic environment of the Spanish/Catalan participants in the experiment reported by Costa and colleagues (2009) fits the definition of this sociolinguistic context, as does the sociolinguistic situation of the Dutch/Frisian participants in our study. Because the bilingual advantage might be found in the interaction between executive control functions responsible for conflict resolution, and monitoring processes, it will probably be most easily detected in a task tapping into these two cognitive functions. Prior and MacWhinney (2010) compared performance by monolingual and fluent bilingual college students in an experiment based on the task-switching paradigm. Instead of an alternating runs paradigm, they used cued task-switching, so that the task not only tapped into conflict resolution, but also into monitoring processes. They found that the bilinguals had reduced switching-costs, which the authors attributed to “transient control processes for selecting between competing tasks, such as activating concurrent task goals and reconfiguring stimulus-response mappings”, and also to “enhanced resistance to proactive interference, a subtype of inhibitory control” (p. 259). Because Prior and MacWhinney’s test proved suitable for detecting differences between bilinguals and monolinguals in the interaction between conflict resolution and monitoring processes, and because our participants’ sociolinguistic situation required exactly such a test, we decided to base our general cognitive test on theirs, after adapting it for our specific age-groups.

With regard to the specific focus of this chapter, namely the the issue to what extent individual differences in language background and language use play a role in task performance, studies reporting on bilingualism and executive control differ in the amount of detail that is used for defining the criteria for bilingualism. In some studies (e.g. Bialystok et al., 2004; Prior & MacWhinney, 2010), the participants speak a variety of different second languages, whereas in others (e.g. Costa et al., 2009), all participants are fluent in the same two languages. Moreover, many reports lack sufficient detail on the participants’ language use, language dominance and the order of acquisition of their languages.

Therefore, various researchers (e.g. Bialystok, 2009) already proposed that research should be done into questions like how much and what type of bilingualism are required to bring about effects on cognitive functioning.

With regard to the question to what extent individual differences, besides language-related factors, within the early-bilingual group play a role in task performance, a large number of studies report an age-related decline in functions depending on executive control (cf. Cepeda, Kramer & Gonzalez de Sather, 2001; Kray & Lindenberger, 2000; Reimers & Maylor, 2005; but see Hobson and Leeds, 2001, for less conclusive evidence). As far as the factor education is concerned, many studies report scores on standardized intelligence tests (for instance Bialystok et al., 2004; 2006) for their participants or, particularly when a study involves college-aged participants, take self-reported SAT-scores as a measure of general cognitive ability (e.g. Prior & MacWhinney, 2010). When the number of years of education is reported, often no distinction is made between educational levels (cf. the study by Kray and Lindenberger, 2000). However, the results of e.g. the study by Meuter and Orr (lecture ISB8 Oslo, 2011) suggest that level of education and language-related factors, such as L2 proficiency, are factors that can affect performance on tasks tapping into executive control and should therefore be logged and taken into account.

3.4. Language background and language use

Because the primary objective of this chapter concerns the issue to what extent individual differences in language background and language use affect task performance within our early bilingual group, we developed a questionnaire containing questions on participants' language background, that was based on the one used by Gullberg and

Indefrey (2003). We based our variables exclusively on participants' use of the languages they learnt in early childhood, i.e. Dutch and Frisian. Details about use of any other languages they subsequently gained proficiency in were logged, but not used for the analyses reported on in this chapter. On the basis of their responses, we developed ordinal scales for the factors language past, current language use, and language preference, as follows:

1 Language past: 5-points scale, with 1 indicating a maximum degree of Frisian language input and use, and 5 indicating a maximum degree of Dutch language input and use.

2 Current language use: 5-points scale, with 1 indicating a maximum degree of Frisian language input and use, and 5 indicating a maximum degree of Dutch language input and use.

3 Language preference: 3 points scale, with 1 indicating a preference for using Frisian, 3 indicating a preference for using Dutch.

Then, as we pointed out, performance on the nonverbal task-switching test seems to tap into the executive function of switching between tasks, or shifting between mental sets (cf. Miyake et al., 2000). We propose that components of the verbal fluency test may to some extent tap into this executive control function as well. It seems logical to assume that this effect is strongest in the case of balanced bilinguals, i.e. bilinguals who use their two languages equally often, because they have the most experience in switching between languages (cf. Bialystok, 2009). Although all our Dutch/Frisian participants learned both their languages before age 6 and have used them on a daily basis ever since, the data recorded in the language background questionnaire show vast individual differences within this group concerning the degree of balance between participants' languages. We therefore created the variable 'language balance', in order to investigate whether this factor plays a role in task performance. We operationalized the variable 'Language Balance' as follows.

For the variables 'language past' and 'current language use', a rating of 1 (maximum degree of Frisian input) or 5 (maximum degree of Dutch input) implies the lowest degree of balance in language use, and therefore leads to rating 1 for 'language balance'. For these variables, a rating of 2 (mostly Frisian input) or 4 (mostly Dutch) implies a higher degree of balance, and is therefore rated as 2 for 'language balance'. A rating of 3 for these variables implies the highest degree of balance and is therefore rated as 3 for 'language balance'. For the variable 'language preference', a rating of 1 (preference for Frisian) or 3 (preference for Dutch) implies a lower degree of balance and is therefore rated with 1 for 'language balance'; a rating of 2 (indicating no preference) implies the highest degree of balance, and is therefore rated as 2 for 'language balance'. These three ratings (i.e., for language past, current language use and language preference) are finally added up; this leads to a variable with a 6-points scale, with a rating of 3 indicating the lowest, and a rating of 8 indicating the highest degree of balance in language use.

3.5. Methodology

3.5.1 Participants

Our experiment involved a group of 52 early bilingual Dutch/Frisian participants; 26 participants belonged to the elderly age-group (mean age 73.2, SD 6.2), and 26 to the middle-aged group (mean age 46, SD 5.7). Participants were assigned to this language group on the basis of the language questionnaire. All participants were fluent in both languages, had acquired them before the age of 7 and used them on a daily basis ever since. Frisian is a minority language spoken in Friesland, a northern province of the Netherlands; it is highly related to Dutch, can be taken as an exam subject in secondary education, and can be studied at the University of Groningen. Task-switching scores of 1 middle-aged participant were discarded because of a physical handicap of both hands, and of 1 elderly participant because of disturbances during testing. Knowledge of and/or proficiency in additional languages were logged but not taken into account as an additional variable. Of the middle-aged group, 10 participants were male, and 16 female; of the elderly group, 13 participants were male and 13 female. All participants were selected according to educational background: they had at least 4 years of secondary education. This is taken as a measure of general cognitive ability.

3.5.2. Procedure, general

All the participants were tested by the same experimenter, in their own homes, during a single experimental session lasting ca. 120 minutes. First, in a test-demo participants were given instructions for the computerized version of the task-switching test. Next, they performed the first part of this task-switching test, to be followed by a verbal fluency test, and then the second part of the task-switching test. Finally, participants carried out the forward and backward versions of the Corsi Blocks test (a spatial working-memory test), and were given a questionnaire to assess their daily language use pattern, their self-rated proficiency and the age at which they acquired their L2 and any other languages they were familiar with. The questionnaire also contained a number of questions on personal and educational background and on factors that might affect cognitive functioning, such as musical practice and computer gaming; originally, it was based on the language background questionnaire by Gullberg and Indefrey (2003).

3.5.3. Task-switching test

This test was based on the experiment reported by Prior and MacWhinney (2010), who adapted their procedure from Rubin and Meiran (2005). The test was presented on a HP laptop computer with a 15.4 inch screen. Participants were seated ca. 60-80 cm from the monitor screen and had to respond to objects presented on the screen, by pressing buttons on a Serial Response Box (produced by Psychological Software Tools Inc., Pittsburgh, PA). Experimental script and data collection were managed by the E-prime computer program for response time measurement. Before the start of the actual test, participants were presented with a set of instructions, also programmed in E-prime, to familiarize them with the general test procedure, and enabling them to ask questions if anything was unclear. The test itself consisted of three blocks, all comprising a set of instructions, to be followed by a number of practice trials and experimental trials. All trials started with the presentation of a fixation cross for 350 ms. followed by a 150 ms. blank screen. Then a task cue appeared on the screen for 650 ms, slightly above the fixation cross. The cue for the color task was a color wheel and the cue for the shape task a black undefinable shape. While the task cue remained on the screen, the target appeared in the center of the screen. Targets were either red or green squares or triangles. Neither task cues nor targets contained any linguistic information, so that the participant’s language could not affect performance on the task. Cue and target remained on the screen until the participant’s response or for a maximum duration of 7 seconds. During practice trials, participants received written feedback on the screen, telling them if their response was correct or not; during experimental trials no feedback was given. After the participant's response a blank screen was presented during 850 ms. which was followed by the start of the next trial.

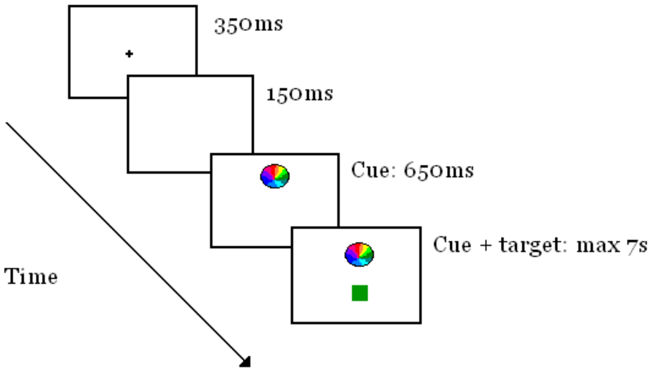

Figure 1: Example of a trial display for color condition (Timmermeister, 2012)

During single-task blocks, the color task had to be performed with the right hand, and the shape task with the left hand. During the color task, participants were instructed to respond to the appearance of a red object with the index finger, and to that of a green object with the middle finger, while ignoring the shape of the object. During the shape task, they had to respond to the appearance of a triangle with the index finger, and to that of a square with the middle finger, while ignoring the color of the object. During mixed-task blocks, the conditions from the single-task blocks were combined and the assignment of task to hand and finger was preserved. This implied that before the presentation of the target either the color circle or the black shape appeared, so that during each individual trial participants either had to focus on the color of the object and ignore its shape, or focus on the shape of the object and ignore its color. The buttons of the Serial Response Box for the color task were labeled with red or green stickers, and the buttons for the shape tasks with stickers showing a triangle or a square in black and white, with similar stickers attached slightly above the buttons. Additionally, the same stickers were pasted slightly below the screen, so that participants would not have to move their gaze from the screen to the Serial Response Box to remember the instructions, thus reducing potential time-delay because of working-memory load.

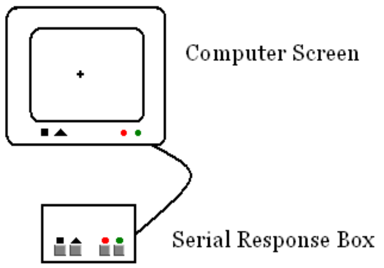

Figure 2: Experiment Set-up Task-Switching Test (Timmermeister, 2012)

The task-switching experiment consisted of two parts. The first part comprised two single-task blocks (first color, then shape), each including 8 practice trials followed by 24 experimental trials, and one mixed-task block including 8 mixed-task practice trials, followed by 48 experimental trials. In both the single and mixed-task blocks, participants could only start with the experimental trials when at least 80% of their responses on the practice trials were correct. Furthermore, 2 dummy trials were added before each group of experimental trials, to reduce effects of time-delay because of starting problems. These dummy trials were not included in the analysis. In the mixed block, half of the experimental trials were switch trials (i.e., participants had to switch from indicating the color, to indicating the shape of the target, or vice versa) and half of them repetition trials (i.e. they had to focus on the same criterion as in the previous trial), ordered in a semi-random design with a maximum of 3 consecutive trials of the same type. After a break, during which participants took part in the verbal fluency test, the second part of the task-switching test was administered. This part started with a mixed-task block, which comprised 8 practice trials and 48 experimental trials, again preceded by 2 dummy trials. The mixed-task block was followed by two single-task blocks, again comprising 8 practice trials, 2 dummy trials and 24 experimental trials, and presented in the opposite order from the first part. Altogether, the experimental blocks in this sandwich design contained 48 switch, 48 repetition trials, and 96 single-task trials (48 color and 48 shape).

3.5.4. Verbal Fluency Test

The test we administered was based on the test from the Delis-Kaplan Executive Function (D-KEF) System (Delis et al., 2001), and comprised three parts.

a. Semantic fluency: in the first part of the test, participants were instructed to name as many words as possible from a semantic category, within 60 seconds. In our experiment, there were two large categories (fruits and vegetables, and jobs) and two small categories (musical instruments, and animals living in or around the water.

b. Phonological fluency: in the second part of the test, participants were instructed to name as many words as possible from a phonological category (i.e., words starting with the same letter), within 60 seconds. We conducted an analysis of CELEX corpora for information about the numbers of exemplars in the testing languages that we used for our participants. For our early bilingual Dutch/Frisian participants, we used the initial letters ‘b’, ‘j’ and ‘r’. This was decided because we needed a high frequency letter (b), a low frequency letter (j) and a moderate frequency letter (r), both in Dutch, the test-language of our early bilingual participants, and in German, the future test language of our monolingual and late bilingual participants. Additionally, our analysis of CELEX corpora showed that these letter categories were equally large in Dutch and German.

c. Category switching: in the final part of the test, participants were instructed to switch between naming exemplars from two semantic categories, again within 60 seconds. First, participants had to name alternatingly musical instruments and animals, and next pieces of clothing and pieces of furniture.

For the components semantic fluency and category switching, the only restriction was that participants should say different words, but that synonyms were allowed. For the phonological task, the following additional restrictions applied: 1 no proper names, of people or places; 2 no numbers; 3 no different word forms or endings, e.g. visiting-visitor; 4 no lengthening of words by compounding.

3.5.5. Corsi Blocks Test

Lastly, we measured participants’ working-memory by means of the Corsi Blocks Test. This is a test of visual spatial working-memory, developed by Milner (1971). It consists of nine identical blocks (3x3x3 cm), irregularly positioned on a wooden board. All blocks have labels with numbers, which are only visible to the experimenter. During the test, the experimenter taps a series of blocks at a rate of 1 block per second. Immediately afterwards, the participant is required to tap the same blocks, in the same order of presentation. The length of block sequences is increased from 2 till 9, with each number (but in a different sequence) being tapped twice. We administered the test both in the forward and the backward versions (here participants had to tap the sequence of blocks in the reverse order from which it was presented). Participants always finished the entire range of sequences, even when recall on two consecutive sequences was not correct. Two measures were used to assess the participants' scores on this task: 1 Span, i.e. the maximum sequence of blocks tapped correctly by the participant; 2 Ratio, i.e. the total number of blocks tapped correctly in the entire test. For our analyses, we used the ratio score on the backward task, because this measure made the finest distinction in performance on the task.

3.6. Results

The mean scores, per age-group, on the verbal fluency test are presented in Table 1.

Table 1 This table shows the verbal fluency scores for the two age-groups (significant effects for age group are signaled as * for p<0.05 and ** for p<0.01).

	Middle-aged N=26 MA 46 (5.7)	Elderly N=26 MA 73.2 (6.2)
Musical instruments	17.4 (5.2)	14.4 (4.4)*
Fruit/vegetables	23.4 (4.7)	21.2 (4.9)
Water animals	16 (5.2)	15.9 (6.1)
Jobs	19.2 (4.8)	17.4 (4.5)
Mean semantic fluency	19 (3.9)	17.2 (3.8)
R	13.7 (4.2)	14 (4.1)
J	9.5 (3.1)	8.7 (4.2)
B	17 (3.8)	17 (4.4)
Mean phonological fluency	13.4 (3.2)	13.3 (3.5)
Category switching, music. instr./animals	17.5 (3.0)	15.5 (3.3)*
Category Switching, clothing/furniture	14.9 (2.8)	13.3 (2.9)*
Mean category-switching	16.2 (2.6)	14.4 (2.7)*

We looked at the correlations of age, education, ‘language use in the past’, ‘current language use’, ‘language balance’ and Working Memory, with phonological fluency, semantic fluency and category switching. The results of this analysis are presented in Table 2.

Table 2 This table shows the correlation coefficients and, between brackets, significance values of the correlations of age, education, 'language use in the past', 'current language use', 'language balance' and Working Memory with phonological fluency, semantic fluency and category switching (significant effects are signaled as * for $p<0.05$ and ** for $p<0.01$).

	Semantic fluency	Phonological fluency	Category switching
Age	r = -.324 (.019*)	r = -0.002 (.988)	r = -.391 (.004**)
Education	r = .239 (.088)	r= .420 (.002**)	r = .204 (.146)
Language use in past	r = .411 (.003**)	r = .253 (.070)	r = .372 (.007**)
Current language use	r = .112 (.428)	r = .140 (.322)	r = .221 (.115)
Language balance	r = .364 (.008**)	r = .204 (.147)	r = .481 (.000**)
Working Memory	r = .441 (.001**)	r = .189 (.180)	r = .518 (.000**)

When we controlled for the factors age and Working Memory, we found that for semantic fluency, the factor 'language use in the past' remained significant at .029 (r= .309) and the factor 'language balance' at p =.047 (r = .283); the factor education became significant at p=.045 (r=.284). For phonological fluency, the factor education remained significant at p = .004 (r = .396). For category switching, the factor 'language use in the past' lost its significance, at p = .093 (r = .240), and the factor 'language balance' remained significant at p = .003 (r = .413). Education became almost significant, at p=.067 (r=.261).

When we only controlled for Working Memory, the factor age lost its significance. The mean scores for the task-switching test, per age-group, are presented in Table 3:

Table 3 This table shows the task-switching results for the two age-groups; RTs in milliseconds/accuracy in %, SDs between brackets (significant effects are signaled as * for p<0.05 and ** for p<0.01).

	Middle-aged N=25; MA 46.3 (5.6)	Elderly N=25; MA 73.6 (6.1)
Single Block, Color /Accuracy	484.8 (92.7)/ 0.98 (0.02)	613.9 (108.3)**/ 0.98 (0.02)
Single Block, Shape/Accuracy	477 (76.5)/ 0.99 (0.01)	632.1 (169.6)**/ 0.99 (0.01)
Mixed Block/Accuracy	694.3 (182.2)/ 0.97 (0.04)	1026.8 (235.7)**/ 0.94 (0.06)*
Mixed, Repetition trials	644.2 (163.6)	953 (217)**
Mixed, Switch trials	746.8 (205.4)	1103.2 (262.7)**
Switch-effect (difference between repeat and switch trials, within mixed blocks)	102.5 (68.5)	150.2 (99.3)
Mix-effect (difference between singleblock trials and repeat trials within mixed blocks)	163.3 (103.8)	331.2 (196.2)**

We looked at the correlations of age, education, ‘language balance’ and Working Memory with the switch and mix effect. The results of this analysis are presented in Table 4.

Table 4 This table shows the correlation coefficients and, between brackets, the significance values of the correlations of age, education, 'language balance' and Working Memory with the switch and the mix effect.

	Switch effect	Mix effect
Age	r = .379 (.007*)	r = .503 (.000**)
Education	r = -.063 (.665)	r = .141 (.329)
Language Balance	r = -.423 (.002**)	r =-.088 (.542)
Working Memory	r = -.444 (.001**)	r = -.524 (.000**)

When we controlled for the factors age and Working Memory, the factor 'language balance' remained significant at p= .024 (r = -.326), implying that the higher the degree of balance between the participants' two languages (Dutch and Frisian), the smaller the switch effect.

Finally, we looked at the correlations of the switch and the mix effect with the components semantic fluency, phonological fluency and category switching of the verbal fluency test. The results are presented in Table 5.

Table 5 This table shows the correlation coefficients and, between brackets, the significance values of the correlations of the switch and mix effect with semantic fluency, phonological fluency and category switching.

	Semantic fluency	Phonological fluency	Category switching
Switch effect	r =-.320 (.024*)	r = -.224 (.117)	r =-.295 (.038*)
Mix effect	r =-.295 (.038*)	r =-.088 (.544)	r = -.243 (.089)

3.7. Discussion and conclusion

The main focus of the present chapter concerned the question to what extent individual differences in language background and language use within the early bilingual group play a role in performance on the cognitive tasks of our study. In our analyses of the results, we also incorporated the factors age, education and working memory, because earlier research suggests that these variables can also affect task performance. Concerning this main research question of our chapter, for phonological fluency we did not find any significant effects for factors denoting language background and language use. The only factor that significantly correlated with the scores on this task was the level of education; when we statistically controlled for Working Memory, this factor retained its significance. The effect of education on semantic fluency and category switching was not significant, but for semantic fluency the effect became significant when we controlled for age and Working Memory. This effect of the factor education on the scores for particularly phonological fluency is in line with previous research findings (cf. Ardila et al., 2000). However, for the task components semantic fluency and category switching we found significant effects for factors relating to language background and language use: both the variables ‘language use in the past’ and ‘language balance’ correlated significantly with task performance. In the first place, this implies that participants who used more Dutch than Frisian during their childhood performed better on these tasks than participants who were relatively more exposed to Frisian, and less to Dutch. This result seems to be in line with earlier findings, i.e. the finding that generally, monolinguals outperform bilinguals in particular on semantic, and not on phonological fluency tasks. The positive correlations we found between the factor ‘language balance’ and semantic fluency and category switching are more challenging to explain. It implies that participants with more balance in the use of their languages perform better on these tasks, as well. We propose that category switching taps heavily into the executive function of shifting between mental sets, and that the higher the degree of language balance of the participants, the more this flexibility has been boosted by their experience in switching between their languages. For semantic fluency switching between categories is not explicitly required. Still, previous studies suggest that in this task component participants’ task strategies involve conscious or subconscious switching between semantic clusters (cf. Unsworth et al., 2011). This evidence that we found for a positive effect of the degree of language balance on the executive function to switch between mental sets in a verbal cognitive task is all the more interesting because our group of participants was very homogeneous. Not only were participants highly fluent in the same two languages, but the vast majority had also spent most of their lives in the same geographical area, i.e. in the north of the Netherlands, and mostly in the province of Friesland.

We also found significant positive correlations between the factors age and Working Memory, operationalized by the scores on the backward Corsi Blocks tests, and the categories semantic fluency and category switching. Because the elderly group scored significantly lower than the middle-aged group on the Corsi Blocks test, we also expected a significant difference between the two age-groups for these task-conditions, but we did not find such a difference for semantic fluency. However, we found that performance on this task was constantly high for participants in their sixties, but rapidly declined after the age of 70. Our finding that this age-related decline in performance disappeared when we controlled for the Corsi Blocks scores, suggests that the effect was related to an age-related decline in Working Memory.

With regard to the scores on the task-switching test, we found a significant negative correlation between the factor language balance and the size of the switch-effect, suggesting that participants with a high degree of balance in their language use suffered lower costs when they had to switch between tasks. We propose that apparently, in accordance with earlier findings (cf. Prior & MacWhinney, 2010), bilingual participants with a high degree of language balance may also perform better on general cognitive tasks that depend on the executive function of shifting between mental sets (cf. Miyake et al., 2000), i.e. the ability to monitor task demands and to switch between tasks.

Last but not least, analyzing interactions between results of the verbal fluency test and the task-switching test, we found significant negative correlations between the scores on semantic fluency and the size of the switch and the mix-effect, and between scores on category switching and the switch effect, suggesting that participants producing a higher number of verbal responses were less troubled by having to switch between tasks. This correlation is in line with the interaction hypothesized by e.g. Bialystok (2008a), according to which bilingualism would simultaneously lead to a reduction in lexical access and a boost in executive control functions. According to this hypothesis, bilinguals "who have developed the most effective control processes (…) would also be the ones least affected by lexical competition and would therefore perform well on (…) fluency tests" (Bialystok et al., 2008a, p. 870). This dissociation that we found between lexical retrieval and executive control in the same participants suggests that the bilingual disadvantage in verbal performance and their advantage in nonverbal cognitive performance could be results of the same underlying mechanism. We propose a mediating role for Working Memory in this effect, as performance on the Corsi Blocks test correlates positively with semantic fluency scores, and negatively with the switch and the mix-effect.

In this chapter, we investigated to what extent individual differences in language background and language use within the early bilingual group play a role in performance on both the general and the verbal cognitive task. In the next chapter, we will compare performance on the general cognitive task between the early bilingual group and the group of functionally monolingual controls.